

Perbandingan Arsitektur Visual Geometry Group dan MobileNet Pada Pengenalan Jenis Kayu

Jimmy Feriawan¹, Daniel Swanjaya²

^{1,2}Teknik Informatika, Fakultas Teknik, Universitas Nusantara PGRI Kediri

E-mail: ¹[*feriawanjimmy@gmail.com](mailto:feriawanjimmy@gmail.com), ²daniel@unpkediri.ac.id

Abstrak – Furnitur merupakan salah satu komoditas strategis bagi ekonomi Indonesia karena furnitur merupakan produk yang bernilai tambah tinggi dan memiliki daya saing global. Produksi furnitur kayu di Indonesia pada tahun mencapai 80% dari keseluruhan produksi furnitur. Untuk mengidentifikasi jenis kayu dibutuhkan seorang pakar kayu. Proses identifikasi kayu bisa ditentukan oleh beberapa faktor seperti warna, tekstur, berat dan masih banyak lagi. Identifikasi kayu bisa dilakukan dengan teknologi Computer Vision karena memiliki keuntungan meningkatkan waktu identifikasi dan mengurangi biaya pada proses Quality Control. Convolutional Neural Network (CNN) merupakan salah satu implementasi metode Deep Learning yang digunakan untuk pemrosesan citra digital. metode CNN dengan arsitektur VGG 16 dan MobileNet dipilih untuk diterapkan pada sistem identifikasi kayu. Setelah dilakukan proses pelatihan dan testing, model arsitektur MobileNet memiliki hasil akurasi senilai 96% dan proses pelatihan juga lebih cepat dibandingkan model arsitektur VGG 16 yang memiliki hasil akurasi senilai 90%.

Kata Kunci — Klasifikasi Kayu, VGG, MobileNet

1. PENDAHULUAN

Furnitur merupakan salah satu komoditas strategis bagi ekonomi Indonesia karena furnitur merupakan produk yang bernilai tambah tinggi dan memiliki daya saing global. Salah satu nilai strategis komoditas furnitur karena industri furnitur merupakan industri padat karya yang menyerap tenaga kerja dalam jumlah yang signifikan. Produksi furnitur kayu di Indonesia pada tahun 2014 mencapai 80% dari total keseluruhan produksi furnitur [1].

Untuk mengidentifikasi jenis kayu dibutuhkan seorang pakar kayu. Proses identifikasi kayu bisa ditentukan oleh beberapa faktor seperti warna, tekstur, berat dan masih banyak lagi. Seiring dengan perkembangan teknologi proses identifikasi kayu bisa dilakukan dengan Computer Vision dengan cara mengidentifikasi citra digital dari tekstur kayu. Identifikasi jenis kayu menggunakan Computer Vision sangat menguntungkan karena bisa meningkatkan waktu identifikasi dan mengurangi biaya pada proses Quality Control.

Computer Vision merupakan ilmu yang dapat membuat mesin dapat melihat, mengidentifikasi dan mengekstrak informasi dari sebuah objek untuk menyelesaikan suatu masalah [2]. Computer Vision memiliki beberapa pendekatan salah satunya yaitu Jaringan Syaraf Tiruan (JST). JST merupakan penerapan cara kerja otak manusia yang diterapkan pada mesin. Pengembangan lebih lanjut dari JST menghasilkan subfield lain yaitu Deep Learning.

Penelitian terdahulu tentang klasifikasi jenis kayu dengan judul “Seleksi Arah Sudut Komputasi dan Fitur GLCM Pada Ekstraksi Citra Kayu Jati, Mahoni, Mindi, dan Sengon” yang dilakukan oleh Stefanus Santosa dan kawan-kawan. Penelitian menghasilkan jumlah fitur yang terpilih yaitu berupa ASM, Kontras, IDM, Entropi dan Korelasi berdasarkan sudut komputasi dari metode GLCM.

Dari tiga jenis sudut komputasi hanya 2 fitur saja yang selalu terseleksi yaitu IDM dan Entropi [3].

Penelitian lainnya tentang klasifikasi jenis kayu yaitu “Klasifikasi Jenis Kayu Menggunakan Back-Propagation Neural Network Berdasarkan Fitur Gray Level Co-occurrence Matrix” yang dilakukan oleh R.A. Pramuendar dan kawan-kawan. Terdapat 4 jenis kayu yaitu Jati, Mahoni, Sengon dan Mindi. Dengan menerapkan fitur gabungan dari 4 sudut komputasi GLCM, input warna grayscale dan RGB maka menghasilkan 80 fitur hasil akurasi tertinggi mencapai 98,27% [4].

Penelitian “Automatic Wood Classification Using a Novel Color Texture Features” oleh Shivashankar S. dan Madhuri R. Kagale. Berdasarkan kombinasi fitur dari metode histogram yaitu Mean, Standard Deviation, Slope, Homogeneity, dan Entropy menghasilkan pada Parquest Database menghasilkan akurasi tertinggi 91,12% dengan jumlah 6 fitur dari kombinasi Standard Deviation dan Slope [5].

Pada penelitian yang dilakukan oleh R.A. Pramuendar dan kawan-kawan [4] data yang digunakan berupa kayu yang memiliki permukaan yang tidak rata dan memiliki bentuk yang berbeda. Penggunaan metode Convolutional Neural Network (CNN) untuk menggantikan metode histogram yang digunakan oleh Shivashankar S. dan Madhuri R. Kagale untuk meningkatkan hasil dari akurasinya.

CNN merupakan salah satu implementasi dari Deep Learning yang digunakan untuk pemrosesan citra digital. CNN menjadi metode populer karena berhasil menempati peringkat pertama sejak tahun 2012 sampai 2015 dalam kategori Image Classification pada sebuah kompetisi global yang diikuti pengembang ternama seperti Google, Facebook, Microsoft, dan lain-lain yang bernama ImageNet Challenges. Dengan classification error terbaik pada tahun 2017 yaitu 0.023% [6].

2. METODE PENELITIAN

2.1 Jaringan Syaraf Tiruan

Jaringan Syaraf Tiruan (JST) menjadi salah satu teknik populer sebagai identifikasi dalam bidang pengenalan pola, sesuai dengan namanya JST merupakan jaringan komputasi yang dibuat untuk mensimulasikan proses pengambilan keputusan dari neuron pada sistem saraf pusat biologis pada manusia maupun hewan. Langkah perhitungan dasar pada sistem biologis manusia umumnya membutuhkan sekitar satu mili detik dibandingkan mesin komputer saat ini yang mampu mencapai nan detik [7].

2.2 Ekstraksi Ciri

Ekstraksi ciri atau fitur adalah proses pengambilan ciri sebuah objek yang dapat menggambarkan karakteristik dari suatu objek. Untuk memproses sebuah citra gambar tidak bisa langsung mengidentifikasi setiap piksel dari keseluruhan gambar yang ada, karena umumnya citra gambar memiliki ukuran yang besar yang akan membebani kinerja dari jaringan syaraf tiruan [8]. Semisal sebuah gambar yang berukuran 10x10 piksel maka weight yg dihasilkan pada layer pertama berjumlah sekitar 10.000 parameter.

2.3 ImageNet Large Scale Visual Recognition Challenges (ILSVRC)

ImageNet Large Scale Visual Recognition Challenges (ILSVRC) merupakan sebuah kompetisi global yang mengevaluasi algoritma untuk deteksi obyek dan klasifikasi gambar pada skala besar yang sudah ada sejak tahun 2010 dan menjadi standar acuan untuk pengenalan obyek skala besar [6]. *Dataset ImageNet* memiliki total 1 juta gambar dengan 1000 kelas dengan masing-masing 50 ribu data untuk *validation dataset* dan 150 ribu data untuk *test set*.

Pada Gambar 1 merupakan contoh tiga kelas *dataset ImageNet* dengan masing-masing kelas berisi banyak jenis obyek.

2.4 Convolutional Neural Network (CNN)

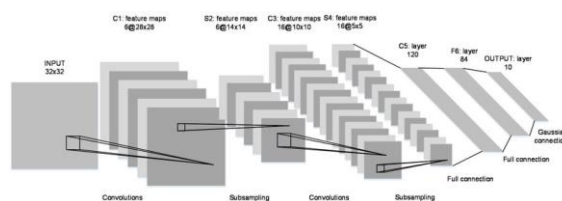
Convolutional Neural Network (CNN) merupakan salah satu metode *Deep Learning*. Arsitektur *CNN* terdiri 5 bagian utama yaitu *input layer*, *convolution layer*, *pooling layer*, *fully-connected layer* dan *output layer* (Gambar 5).

Tujuan dari proses pelatihan adalah melatih model jaringan syaraf tiruan untuk meminimalkan error hasil prediksi model dengan data asli. *Input layer* adalah vektor dari gambar dataset. *Convolution layer* merupakan operasi konvolusi antara 2 vektor. Pada persamaan 1 merupakan konvolusi dua buah fungsi dimana $g(x)$ disebut sebagai kernel konvolusi (filter) yang akan dioperasikan secara bergeser pada vektor $F(x)$.

$$h(x) = F(x).g(x) = \int F(a).g(x - a) \dots\dots\dots (1)$$



Gambar 1. Contoh *dataset ImageNet*



Gambar 2. Arsitektur CNN LeNet 5 angka tulisan tangan [8].

Tabel 1. Konfigurasi arsitektur VGG

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

2.5 Visual Geometry Group (VGG)

Visual Geometry Group (VGG) merupakan salah satu model arsitektur *CNN* yang diusul oleh K. Simonyan dan A. Zisserman dari Universitas Oxford pada kompetisi ILSVRC-2014.

Konfigurasi arsitektur VGG pada tabel 1 dibagi menjadi 5 bagian berdasarkan jumlah kedalaman *layer*. Arsitektur VGG16 berarti memiliki 16 kedalaman *layer* pada konfigurasinya [9].

“Pengenalan Golongan Jenis Kendaraan Bermotor pada Ruas Jalan Tol Menggunakan CNN” yang dilakukan oleh Ricky Herwanto dan kawan-kawan. Penelitian ini menggunakan arsitektur VGG16 untuk mengidentifikasi 5 jenis kendaraan pada jalan tol menghasilkan akurasi sebesar 90,76% dan *f-score* sebesar 73,53% [10].

2.6 MobileNet

Tren umum pembuatan model *Convolutional Neural Network (CNN)* adalah semakin banyak jumlah layer untuk mencapai hasil akurasi yang lebih tinggi, tren ini juga dibuktikan dengan hasil analisa dari ILSVRC dari tahun 2010 sampai 2015 (Gambar 3). Jumlah layer hamper selalu mengalami peningkatan bahkan pada model ResNet jumlah layer mencapai 100 lebih.

Penerepan jumlah layer yang semakin banyak ini tidak serta merta membuat model menjadi lebih efisien karena berkenaan dengan ukuran dan kecepatan karena pada penerapan di dunia nyata seperti pada bidang robotik, self-driving car dan augmented reality diperlukan waktu seefisien mungkin karena keterbatasan kemampuan komputasi perangkat. MobileNet merupakan sebuah model arsitektur CNN yang didesain efisien dengan 2 set *hyper-parameters* untuk membangun model yang sangat kecil dan latensi rendah yang akan dengan mudah diimplementasikan sesuai kebutuhan *mobile* dan *embedded applications*. MobileNet dibuat berdasarkan *depthwise separable convolutions* untuk mengurangi komputasi di *layer* awal [11].

“Aplikasi Convolution Neural Network untuk mendeteksi jenis-jenis sampah” yang dilakukan oleh Stephen dan kawan-kawan. Penelitian ini menggunakan metode CNN dengan membandingkan beberapa arsitektur yaitu VGG 16, MobileNet V1, Inception V3 dan ResNet 50. Resnet 50 menjadi model terbaik dengan nilai akurasi yang didapatkan dari *train* adalah 78% dan 90%. Sedangkan nilai akurasi dan *validation* sebesar 74% dan 80%. Resnet 50 bisa di implementasikan apabila server atau komputer memiliki spesifikasi yang, untuk perangkat dengan spesifikasi rendah seperti kontroler Raspberry Pi maka bisa menggunakan model MobileNet [12].

2.7 Analisa Sistem

Desain arsitektur yang akan digunakan adalah VGG16 dan ResNet50. Untuk layer pooling memiliki perbedaan yaitu *Max Pooling* pada VGG16 dan *Average Pooling* pada Resnet50. VGG16 berarti menggunakan kedalaman sebanyak 16 *layer* pada model JST yang digunakan dan *MobileNet*.

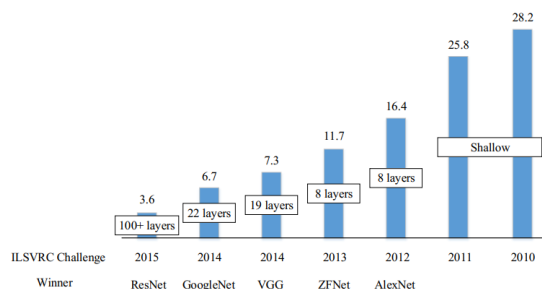
Total *training parameters* untuk model arsitektur VGG16 adalah 27.829.573 dan MobileNet 5.855.429 adalah. implementasi *depthwise separable convolutions* pada MobileNet berdampak berkurang nya *training parameters*.

2.8 Desain Sistem

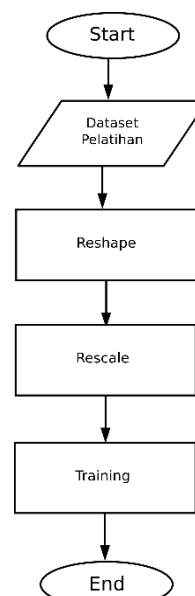
Metode *Convolutional Neural Network (CNN)* tidak memerlukan langkah *pre-processing*, persiapan untuk data sebelum proses pelatihan berupa *reshape* yaitu mengubah dimensi data menjadi *array* 3 dimensi 224x224x3 yang berarti setiap dimensi *array* berisi nilai warna *RGB* gambar dengan rentang 0-255 kemudian nilai tersebut dilakukan proses *rescale* menjadi nilai dengan rentang 0-1 (Gambar 4).

Model algoritma *Convolutional Neural Network (CNN)* dibuat menggunakan bahasa pemrograman python dan dengan library Tensorflow dan Keras. Jenis perangkat yang digunakan untuk melakukan *training* algoritma adalah laptop Acer E5-473G-367E dengan spesifikasi:

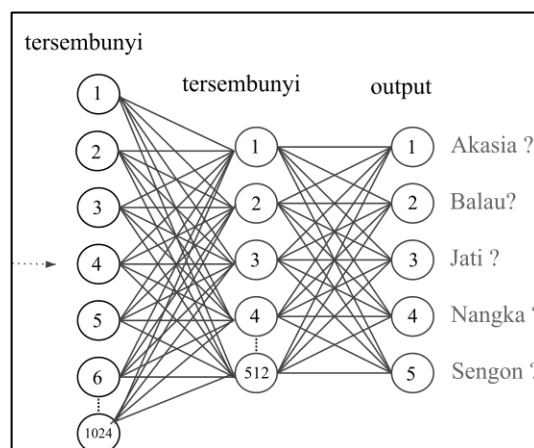
- Prosesor: Intel Core i3-5005U (2.0GHz)
- RAM : 8 GB DDR3L



Gambar 3. arsitektur model CNN pemenang ILSVRC [6]



Gambar 4. Desain proses pelatihan



Gambar 5. Desain arsitektur *fully-connected layer*.

2.9 Desain Arsitektur

Desain arsitektur yang digunakan pada penelitian ini adalah *VGG16* dengan kedalaman layer 16 dan *MobileNet* dengan kedalaman layer 28. Max pooling digunakan pada *VGG 16* dan average pooling pada *MobileNet*. Input yang yang digunakan adalah gambar dari dengan mode *RGB* kemudian gambar di-*resize* menjadi ukuran 224x224 piksel dan menghasilkan *array* 3 dimensi. Model *VGG16* dan *MobileNet* di inisiasi dengan model pra-pelatihan *imagenet*.

Gambar 5 merupakan desain arsitektur pada –*ully-connected layer* untuk *VGG 16* dan *MobileNet*. Pada gambar tersebut memiliki 2 *hidden layer* dengan masing-masing 1024 dan 512 neuron. Pada *output layer* terdapat 5 neuron sesuai dengan kelas pada dataset.



Gambar 6. Sampel data pelatihan

3. HASIL DAN PEMBAHASAN

3.1 Pengambilan Data

Penelitian menggunakan *Convolutional Neural Network (CNN)* bertujuan untuk melakukan perbandingan hasil klasifikasi moel arsitektur *VGG 16* dan *MobileNet* terhadap data citra digital pada 5 jenis kayu yang sering digunakan untuk bahan furnitur di rumah yaitu Akasia, Balau, Nangka, Jati dan Sengon. *Dataset* pada Gambar 6 diambil menggunakan ponsel Redmi Note 8 dengan memanfaatkan sensor makro sebesar 2MP dengan ukuran data masing-masing antara 800 hingga 1200 *KiloBytes (KB)*. Dimensi gambar 1200x1200 piksel dilakukan *resize* ke dimensi 224x224 piksel sehingga ukuran data menjadi sekitar 10 hingga 15 KB.

3.2 Hasil Pelatihan

Pada proses pelatihan masing-masing arsitektur dilakukan 3 kali percobaan dengan nilai epoch pada sesuai dengan Tabel 2 dan Tabel 3.

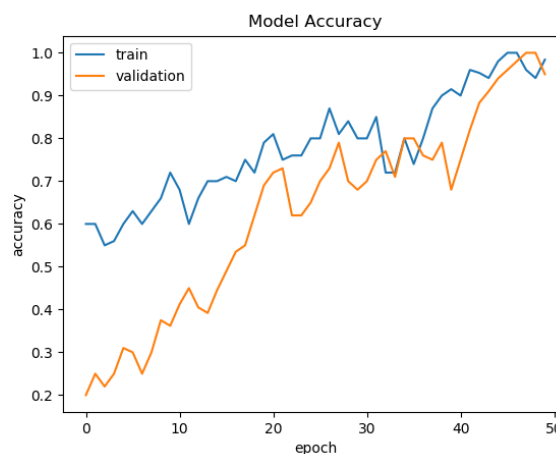
Proses pelatihahn terbaik untuk arsitektur *VGG 16* dan *MobileNet* adalah pada percobaan ketiga dengan masing-masing 50 *epochs* dan 25 *epochs*. Gambar 7 dan Gambar 8 merupakan riwayat pelatihan terbaik untuk masing-masing model arsitektur.

3.3 Hasil Test

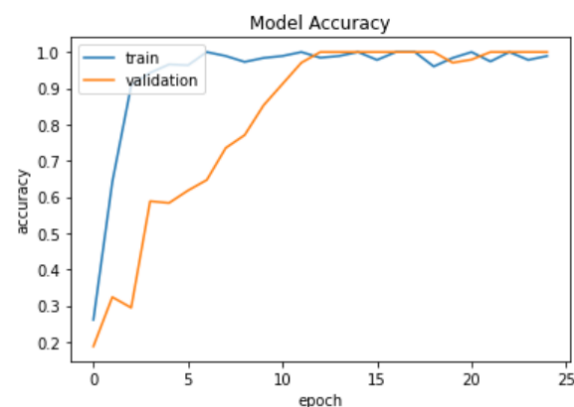
Akurasi Model *JST* diuji melalui proses identifikasi model terhadap data baru yang tidak sama dengan data pelatihan. masing-masing kelas diuji dengan 10 data gambar.

$$Accuracy = \frac{\text{jumlah prediksi benar}}{\text{jumlah data tes}} * 100\% \dots (2)$$

Tabel 4 dan Tabel 5 adalah hasil tes klasifikasi model untuk data test, kolom A,B,J,N dan S adalah inisial dari masing-masing jenis kayu. Kolom T dan F merupakan presentase klasifikasi benar dan salah. Pada masing-masing jenis kayu digunakan 10 data citra untuk pengujian.



Gambar 7. Proses training VGG 16



Gambar 8. Proses training MobileNet

Tabel 2. Hasil training VGG 16

Epoch	Akurasi training (%)	Cross validation (%)	Waktu (menit)
10	74,2	66,7	27
25	84,7	74	38
50	100	94	57

Tabel 3. Hasil Training MobileNet

epoch	Akurasi training (%)	Cross validation (%)	Waktu (menit)
10	96,8	90	14
15	100	96	18
25	100	100	24

Tabel 4. Hasil pengujian VGG 16

kategori	Klasifikasi					Presentase	
	A	B	J	N	S	T	F
Akasia	10	0	0	0	0	100	0
Balau	0	8	0	0	2	80	20
Jati	0	0	8	2	0	80	20
Nangka	0	0	0	10	0	100	0
Sengon	0	1	0	0	9	90	10

Tabel 5. Hasil Pengujian MobileNet

kategori	Klasifikasi					Presentase	
	A	B	J	N	S	T	F
Akasia	10	0	0	0	0	100	0
Balau	0	9	0	0	1	90	10
Jati	0	0	9	1	0	90	10
Nangka	0	0	0	10	0	100	0
Sengon	0	0	0	0	10	100	0

Berdasarkan tabel tersebut kedua model VGG 16 berhasil mengklasifikasikan kayu jenis Akasia dan nangka sedangkan untuk kayu jenis balau dan jati berhasil diklasifikasikan sebanyak 8 citra, kayu sengon sebanyak 9 citra. Untuk model MobileNet berhasil mengklasifikasikan jenis kayu Akasia, Nangka dan Sengon, sedangkan untuk kayu jenis balau dan jati berhasil diklasifikasikan sebanyak 9 citra.

Berdasarkan perhitungan pada persamaan 2 maka akurasi yang dihasilkan model VGG 16 adalah 90% sedangkan untuk model MobileNet adalah 96%.

4. SIMPULAN

Setelah dilakukan analisa sistem, desain sistem, pengimplementasian dan pengujian terhadap sistem yang telah dibuat dengan jumlah dataset 350 yang dibagi untuk proses pelatihan 200, validation 50 dan test 50 dapat ditarik kesimpulan sebagai berikut:

- 1) Performa dari model Convolutional Neural Network (CNN) dengan arsitektur MobileNet memiliki hasil akurasi yang lebih baik senilai 96% dibandingkan VGG 16 dengan akurasi 90%
- 2) Selain memiliki akurasi yang lebih baik arsitektur MobileNet juga membutuhkan waktu lebih cepat pada proses pelatihan dan memiliki ukuran *weight* yang lebih kecil.
- 3) VGG 16 dan MobileNet memiliki error klasifikasi pada jenis kayu yang sama yaitu kayu Balau dan Jati.

5. SARAN

Saran untuk pengembangan kedepannya adalah:

- 1) Meningkatkan jumlah dataset yang lebih bervariasi karena hasil pengujian dan hasil pelatihan yang berbeda.
- 2) Analisa lebih lanjut dan perbaikan pada data jenis kayu Balau dan Jati karena masih terdapat error klasifikasi.

- 3) Menguji konfigurasi lain pada CNN agar dapat meningkatkan akurasi *testing* seperti penerapan *Regularization* pada proses pelatihan.

DAFTAR PUSTAKA

- [1] Salim, Zamroni., Ernawati Munadi. Info Komoditi Furnitur, https://bppp.kemendag.go.id/2017/10/Isi_BRIK_FU_RNITUR.pdf diakses pada 10 Maret 2020.
- [2] Gasim, 2014. Metode Identifikasi Kayu Berdasarkan Model Blok Citra Mikroskopis Penampang Lintang., Program Studi S3 Ilmu Komputer, Jurusan Ilmu Komputer dan Elektronika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Gadjah Mada. Yogyakarta.
- [3] Santosa Stefanus, Martono., Marchus Budi Utomo., Basuki Setiyo Budi. 2018. Seleksi Arah ESudut Komputasi dan Fitur GLCM Pada Ekstraksi Citra Kayu Jati, Mahoni, Minda dan Sengon. Jurnal Wahana Teknik Sipil. No. 23 Vol. 2. 77-87.
- [4] Pramuendar R.A., D.P. Prabowo., D. Pergiwati., K Latifa. 2017. Klasifikasi Jenis Kayu Menggunakan Back-Propagation Neural Network Berdasarkan Fitur Gray Level Co-Occurrence Matrix. Makalah. Dalam: Science and Engineering National Seminar 3 di Universitas PGRI Semarang, 19 Agustus.
- [5] S., Shivashankar, dan Maduri R. Kagale. 2018. *Automatic Wood Classification using a Novel Color Texture Features. International Journal of Applications*, No. 27. Vol 180. 0975-8887, <https://www.ijcaonline.org/archives/volume180/number27/29147-2018916648> diakses pada tanggal 10 Desember 2019.
- [6] Priddy Russakovsky, Olga & Deng, J. & Su, Hao & Krause, J. & Satheesh, Sanjeev & Ma, S. Huang, Z. & Karpathy, A. & Khosla, A. & Bernstein, M.. 2015. *ImageNet Large Scale Visual Recognition Challenge*. Int. J. Comput. Vis. 115(2015):. 1-42.
- [7] Graupe, Daniel. 2019. *Principles of Artificial Neural Network: Basic Design to Deep Learning*. No. 4. Vol. 8.
- [8] LeCun, Yann., Patrick Haffner., Leon Botton., Yoshua Bengio. 1998. Object Recognition with Gradient-Based Learning. *Proceedings of the IEEE*, 86(5):755-824.
- [9] Simonyan, Karen., Andrew Zisserman. *Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv:1409.1556*, 2015. <https://arxiv.org/pdf/1409.1556>. Diakses pada 1 Maret 2020
- [10] Herwanto, Ricky., Kartika Gunadi, Endang Setyati. 2020. Pengenalan Golongan Jenis Kendaraan Bermotor pada Ruas Jalan Tol Menggunakan CNN. *Jurnal Infra*. No. 1 Vol. 8. 196-202.
- [11] G. Howard, Andrew., Weijun Wang., Menglong Zhu., Tobias Weyand., Bo Chen., Dmitry Kalenichenko., Hartwig Adam. *MobileNets: Efficient Convolutional Neural Networks for mobile Vision Applications*. arXiv:1704.04861, 2017. <https://arxiv.org/pdf/1704.04861>. Diakses pada 1 Maret 2020.
- [12] Stephen, Raymond., Handri Santoso. 2019. Aplikasi Convolution Neural Network Untuk Mendeteksi Jenis-jenis Sampah. *Jurnal Sistem Informasi dan Telematika*. No. 2. Vol. 10. 122-132.

[Halaman ini Sengaja Dikosongkan]